The societal impact of research undertaken by Hong Kong universities:
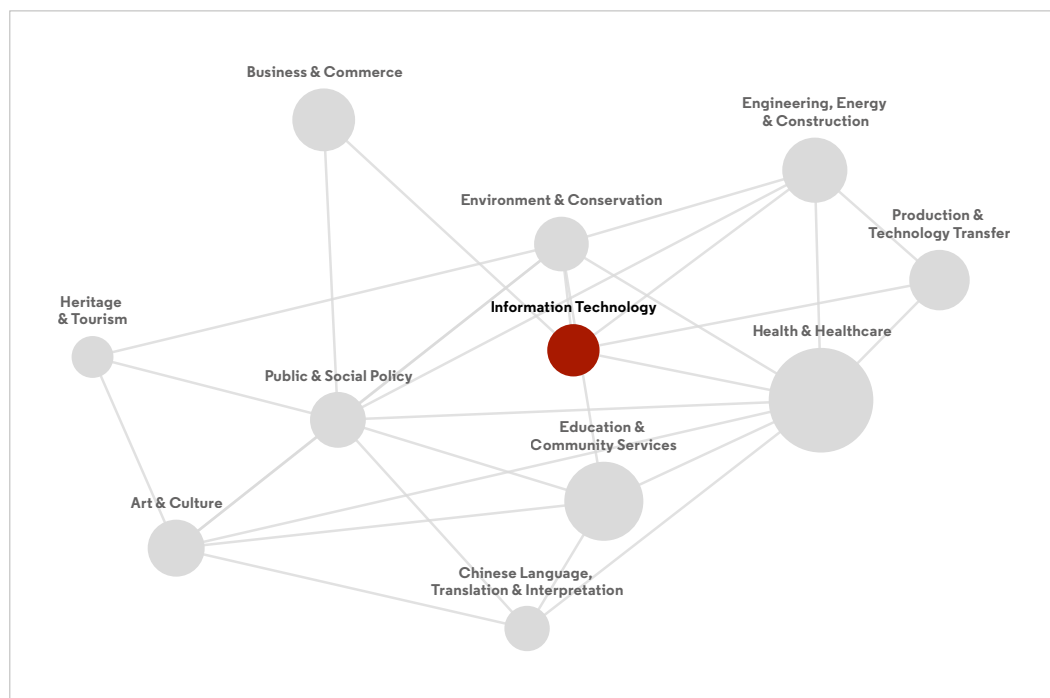
Information Technology

A synthesis of the RAE 2020 impact case studies

Business & Commerce

Engineering, Energy & Construction

Environment & Conservation

Production & Technology Transfer

Heritage & Tourism

Information Technology

Health & Healthcare

Public & Social Policy

Education & Community Services

Art & Culture

Chinese Language, Translation & Interpretation

Partnered with:

ELECTRIC DATA solutions

Different Angles

This report is part of a series of outputs that examines the impact of research arising from eight universities based in Hong Kong and funded by the University Grants Committee (UGC). The report focuses on the Impact Case Studies (ICS) produced by the UGC-funded universities as part of their response to a Research Assessment Exercise (RAE) in 2020. The overarching report - *The impact of research undertaken by universities in Hong Kong: A synthesis of the RAE 2020 impact case studies* – is accompanied by 11 thematic reports that examine the nature of research impact in different areas, ranging from Arts & Culture to Health & Healthcare. The 342 impact case studies that are analysed through this body of work are also available on a searchable database that is posted on the UGC's website.

The Information Technology (IT) cluster contains 24 impact case studies (ICS) from one primary topic identified in the topic modelling.[1] The IT cluster represents 7% (i.e. 24/342) of ICS submitted to RAE 2020.

The impact wheel (Figure 1) shows that these 24 case studies came from seven Units of Assessment (UoAs). For example, for UoA 13 (computer studies/science) eight out of 15 (=53%) ICS submitted to the UoA were in the IT cluster. The six other UoAs were: UoA 11 (mathematics and statistics); UoA 12 (electrical & electronic engineering); UoA 18 (planning and surveying); UoA 22 (business); UoA 29 (communications & media studies); and UoA 38 (visual arts, design, creative media, other creative arts and creative writing).

Figure 1: Impact wheel for the Information Technology cluster (n=24)
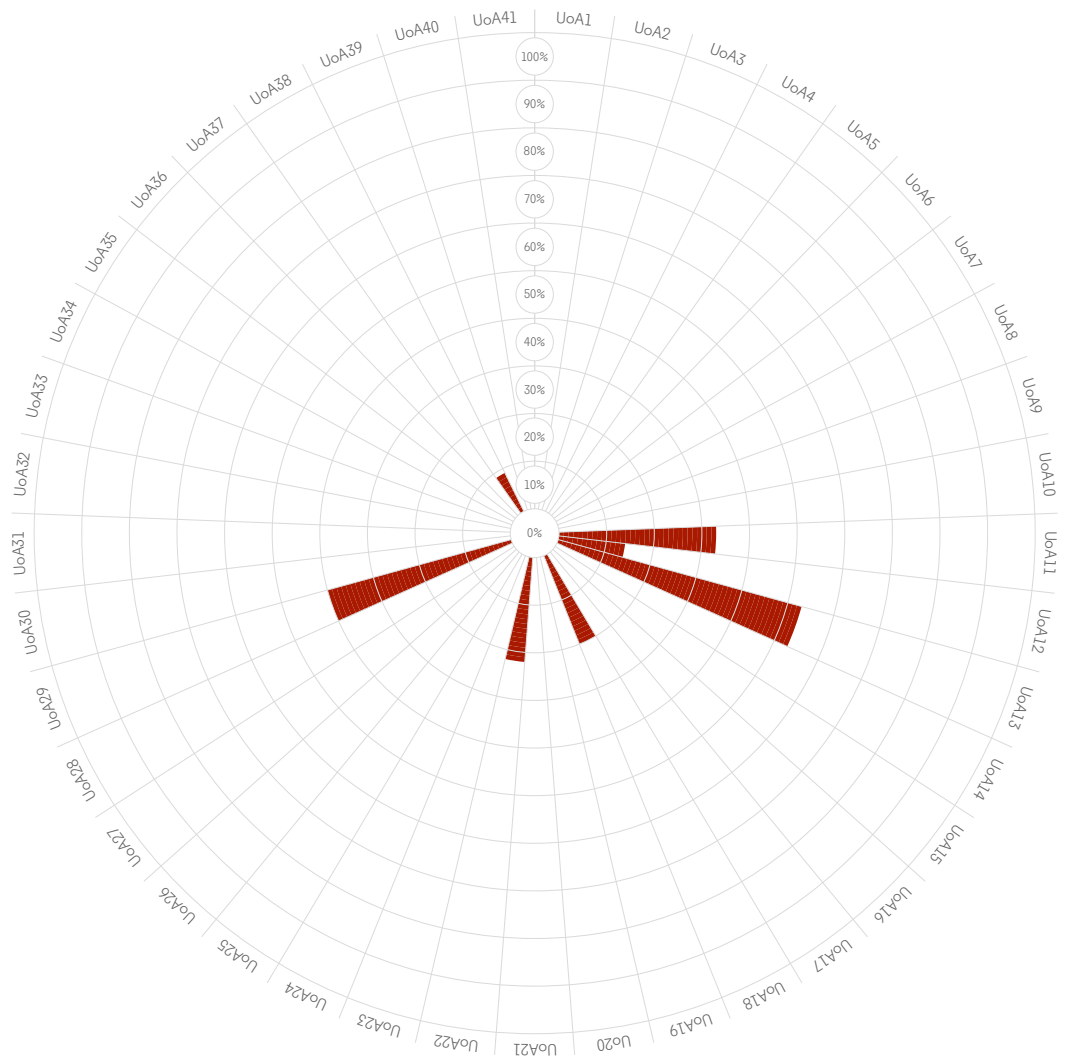
The impact of Hong Kong universities' research: **Information Technology**

Table A shows the most salient features of the case studies in terms of beneficiaries, location, type of impact and time lag. It gives the percentage of case studies in this cluster that were tagged with sub-codes under these code headings, as well as the percentage of case studies tagged with those sub-codes in the entire sample of 342.

The 24 case studies in this cluster impacted beneficiaries in three key sectors under the classification of the Hong Kong Standard Industrial Classification: 58% benefited the Information and communications sector, 29% Professional, scientific and technical activities and 17% Transportation, storage, postal and courier services. The key socioeconomic group was citizens/communities, with 50% of case studies relating to this group, followed by university students (8%). The primary decision taker groups that were involved in the impact was the private sector (63%), government departments/agencies (29%) and NGOs/third sector (21%). Excluding Hong Kong (54%), the Greater Bay Area (4%) and Mainland China (13%), these case studies primarily had an impact in the United States (29%) and the United Kingdom (17%). The most salient type of impact was service or product in regular use (75% of case studies in this cluster), indicating a high degree of application and commercialisation. This was followed by informing procedures, practice or protocol (71%), improving workflows (54%) and generating profit/revenue or cost savings (38%). Interestingly, on average, the research in this cluster was started in 2011, compared to 2006 for the whole sample. This suggests a short time lag between the start of the research and the relevant impact. The median publication date for this cluster was 2015, which mirrored the publication date for the whole sample.

The majority of ICS could be grouped into four main subthemes: Computing/mobile technology; Genetic and health technology; Transport technology; and, other areas of technology.

Table A: Some salient features of research impact identified in the Information Technology cluster (n = 24)

Beneficiaries of impact (top mentions)	% of <u>cluster</u> impact case studies	% of <u>all</u> impact case studies
Hong Kong Standard Industrial Classification		
Information and communications	58%	11%
Professional, scientific and technical activities	29%	13%
Transportation, storage, postal and courier services	17%	5%
Sociodemographic group		
Citizens/communities	50%	17%
University students	8%	6%
Decision taker group		
Private sector	63%	30%
Government departments/agencies	29%	31%
NGOs/third sector	21%	17%
Location of impact	**% of <u>cluster</u> impact case studies**	**% of <u>all</u> impact case studies**
Hong Kong	54%	75%
Greater Bay Area (excluding Hong Kong)	4%	3%
Mainland China (excluding Hong Kong and GBA)	13%	12%
United States	29%	32%
United Kingdom	17%	17%
Type of impact (top mentions)	**% of <u>cluster</u> impact case studies**	**% of <u>all</u> impact case studies**
Service or product in regular use	75%	28%
Inform procedure, practice or protocol	71%	52%
Improve workflows	54%	15%
Generate profit/revenue or cost savings	38%	18%
Elapsed time	**<u>Cluster</u>**	**<u>All</u>**
Median year of research commencement	2011	2006
Median year of publication date	2015	2015

Computing/mobile technology

One area of impact in computing and mobile technology relates to security. For example, City University of Hong Kong researchers developed new techniques and automated software to enable Scalable Security Testing and Code Analysis, which directly led to the discovery of multiple high-impact security vulnerabilities/attacks in single-sign-on and mobile payment systems. The exposure of these vulnerabilities successfully eliminated critical security/privacy risks for billions of stakeholders, and the work has become a key reference in global open access standards. Another ICS, this time from The Hong Kong Polytechnic University researchers created novel solutions to detect and analyse malware targeting the Android system and mobile apps. They developed practical systems to facilitate security researchers and professionals as well resources that allow the public to scrutinize mobile malware and legitimate apps. In another example, The University of Hong Kong's Department of Computer Science addressed the security needs of industry and wider society by creating a practical protocol for extending legal professional privilege to digital evidence and developing a product "SHIELD" (Smart Hacking and Intrusion Entrapment with Lawful Detection) to assist traditional industries to protect themselves from cyberattacks. One ICS details The Hong Kong Polytechnic University's work on mobile system security and privacy, blockchain infrastructure and security, and privacy enhancing technology, which has successfully been deployed in a number of commercial settings. Amongst other innovations, the team has led the field in protecting user privacy in popular cryptocurrencies, and this technology has been adopted by commercial platforms (CryptoNote, Monero and ZCash) with a total market cap of US$40 billion. In a related but distinct ICS, City University of Hong Kong developed software to combat online surveillance (which was downloaded 1.3 million times) and framed the public debate on online surveillance and violations of personal privacy through public exhibitions and educational materials.

Another example of impact in computing and mobile technology relates to reliability of software. The ADORE framework, created by The Chinese University of Hong Kong researchers, makes a substantial contribution to reliability in software operation and maintenance. The tools and datasets of this framework have three applications: a) assessing the quality of web services, b) automatic detection and identification of software reliability issues, c) extracting actionable insights from user feedback at app stores. The quality of web service tools and datasets have been taken up by more than 370 research institutes including Microsoft, IBM, Amazon, Carnegie Mellon University, Imperial College London and cited over 2,000 times, while the automatic detection datasets were downloaded more than 4,000 times by over 200 organizations globally such as ETH Zurich, Microsoft, HSBC. The extraction of user feedback insights tool has been successfully deployed at Tencent, one of the largest IT companies in China, and Huawei, a world leading telecommunications company. Together, this research has improved the usability of modern software, such as Web services, software systems and mobile applications. A similar example of this type of impact relates to the work of City University of Hong Kong, which developed new methods of improving the capacity of mobile storage that has been successful deployed at Tencent and Huawei.

Genetic and health technology

Another area of impact is genetic and health technology. Researchers at The University of Hong Kong developed software that can test a range of kinship scenarios (e.g., missing person cases, incest cases, inheritance dispute cases) to resolve various complex familial relationships. This has provided a simple way of assessing the weight of DNA evidence, which has been deployed by forensic and paternity laboratories globally to evaluate thousands of cases, with strongly positive feedback. Strikingly, the tool was successfully used to identify three badly burnt bodies in the Hot-Air Balloon Burst Tragedy in Egypt. In another ICS, researchers from the Department of Computer Science at The University of Hong Kong reduced the time required and improved the accuracy of DNA next-generation-sequencing data, which was extended to be used in medical diagnosis. This technology was taken up by the Department of Health in 2015 to diagnose genetic disorders, improving the number of cases from around 100 per year to 1,900 cases in the first 1.5 years. A spin-off company was founded in 2014, which established the first bioinformatics cloud in China (BGI Online) and garnered HK$60M (cUS$7.6m) of investment/revenue. A third ICS relates to large-scale genetic data analysis. Researchers at The Hong Kong University for Science and Technology established novel methods that were used in direct-to-consumer DNA ancestry testing platform and personalized healthcare testing provider WeGene. Using these methods WeGene grew from 8,000 customers in 2016 to about 300,000 customers in July 2019. The company's work attracted global media attention on Asian heritage and health issues and had a direct benefit for global data analysis practitioners by creating a standard computer toolkit for handling genetic data. A final ICS details how the Advance Intelligent Information Systems research centre (ATIS) of City University of Hong Kong created a comprehensive AI-enabled framework for healthcare providers to improve efficiency, efficacy, and flexibility of care. By adding smart medical device and real-time diagnosing algorithms to healthcare systems for elderly people, these advances led to improved routine tasks of medical personnel and caregivers, potentially leading to improved health outcomes. In addition, medical device developers expanded their businesses, and policymakers gained insights how to further develop AI-healthcare quality.

Transport technology

Several case studies related to the area of transport and location-based technology. One of the most impactful ICS read across all the themes was the development of an algorithm to support flight control systems for drones. Whilst a postgraduate student, under supervision of one of HKUST professor's, the researcher conducted initial research on autonomous flight control system for unmanned aerial vehicles (UAVs). He developed an unmanned miniature helicopter that achieved the world's first flight on Mount Everest, almost reaching the summit. The team also used the exponential coordinate theory to improve the flight controller. This directly led Wang to commercialize drones through the founding of DJI Technology. DJI integrated the flight control system and navigation technology developed by HKUST into its products, resulting in improved performance and enabling it to become the world leading manufacturer of UAVs. DJI's drones have subsequently been at the forefront of camera drone technology by creating safe, reliable, easy-to-use high-quality aerial imaging systems for creative professionals around the world. In 2017, DJI reported revenues of CNY18 billion (cUS$2.8 billion). Another interesting example is the work of The Hong Kong Polytechnic University, which developed big data analytics for highly efficient transportation, logistics and airport resource management as well as deep learning approaches for enhanced image recognition. This research was operationalised by Hong Kong R&D Centre for Logistics and Supply Chain Management Enabling Technologies (LSCM) and leading logistics organisations, including the Hong Kong International Airport Authority (HKIA) and GoGoVan, and two of the largest technology companies, Huawei and Alibaba.

The research has won HK$20.58m (cUS$2.6m) of funding, and has led to 7 patents, 11 sponsorships, and use within 8 systems with industrial and government partners. In another ICS, City University of Hong Kong researchers were commissioned by the HKIA to develop two systems to improve operations and on-time turnaround performance. These were successfully delivered, bringing up baggage handling on-time performance from 70% to 90+%, and further work valued at HK$27.9m (cUS$3.6m) was commissioned to create more functions for all remaining operations in the airfield. Crucially, an additional benefit was airport management adapting to a management-by-data approach for the first time. An additional example comes from The Hong Kong Polytechnic University team, which developed pathfinding network research into a detailed route guidance website for any user-chosen origin-destination based on multiple criteria. This has provided an important public service, as previously this was only offered by specific companies making integrating multiple criteria around topography, modes, fares, and time very difficult. In a final ICS in this subtheme, a The Hong Kong University for Science and Technology researcher developed lightweight and robust location-based technologies for commercial use. His work has made three key innovations: a) technology for providing indoor location information, b) a low-cost streaming technology that accommodates user mobility, and c) a wifi system that can work in complex environments. These technologies have been turned into commercial applications benefiting millions of users, generated licensing income to HKUST in excess of HK$2m (cUS$0.25m) and generated more than HK$100m (cUS$12.7m) of investment, revenue and company valuation. Having won several prestigious awards, this research has been commended for its innovation, commercial impact and technological leadership.

Other areas of technology

Two other case studies develop processes for gathering insights from sponsored search advertising and TV audience measurement to allow industry to gain strategic insights and increase revenue. Two case studies relate to technological innovations in finance. One describes work at Hong Kong Baptist University, which provided cutting-edge algorithms for constructing scrambled digital sequences used in MATLAB, a commercial mathematical software system with 3 million users, and by Koi Investment Partners International (Koi), an independent quantitative asset management company based in Hong Kong. For Koi Investment Partners, the efficiency of the evaluation and the profitability of trading were increased by at least 10%. Another ICS details City University of Hong Kong's work developing an estimator for assessing the volatility of a stock price for professional traders. This technology has been acclaimed by traders and implemented in many professional websites and is said to be 14 times more efficient than other methods.

The characteristics and translation of the underpinning research

Table B provides the salient features of the underpinning research. It provides bibliometrics as well as information on the impetus for the research and mechanisms/channels of dissemination.

94 outputs from this cluster are indexed on the Web of Science, which have a mean citation score of 8.49, which is very high compared to the whole sample. The median citation score is 1.83, which is higher than the median of 1.59 for all case studies. Key international collaborators included the United States (17%), Singapore (6%), and the UK (4%). 38% of the research was in response to demand for better innovation and technology and 17% was commissioned.

The research often followed on from previous work (38%) and was driven by external advances creating new questions (29%). The main forms of co-production and collaboration were academic and industry partnership (46%) and academic and public sector partnership (17%). The researchers were involved in the impact through collaboration with the sector in 54% of cases, and through co-producing new technology or products (38%). The research findings were disseminated primarily through open-source data websites and databases (38%) and through non-academic presentations (29%). Many ideas and products arising from the research were given prizes and awards (54%) and were formally registered (8%).

Table B: Some salient features of the underpinning research identified in the Information Technology cluster (n = 24)

Analysis of underpinning research	Cluster impact case studies	All impact case studies
Bibliometrics indicators		
Number of outputs indexed on Web of Science	94	1445
Mean citation score	8.49	4.45
Median citation score	1.83	1.59
Collaborators location (top mentions, excluding China)		
USA	17%	18%
Singapore	6%	3%
UK	4%	8%

Impetus for research (top two mentions)	% of cluster impact case studies	% of all impact case studies
Pull factors		
Demand for better innovation & technology	38%	11%
Commissioned	17%	16%
Push factors		
Follow on from research team's previous work	38%	12%
External advances creating new questions	29%	6%

Mechanisms/channels of impact (top mentions)	% of cluster impact case studies	% of all impact case studies
Coproduction & collaboration		
Academic - industry partnership	46%	23%
Academic - public sector partnership	17%	17%
Researcher involvement		
Collaboration with sector	54%	20%
Co-produced new technology/product	38%	9%
Dissemination of research findings		
Open-source data website and databases	38%	15%
Non-academic presentation (incl public lecture)	29%	36%
Codification of impact eg prizes, patents etc.		
Prizes and awards	54%	33%
Idea/product/invention registered	8%	2%

The alluvial diagram in Figure 2 links the underpinning research (as classified by discipline using the 23 Web of Science, Essential Science Indicators (ESI), journal categories) to the 11 clusters identified through the topic modelling and the 13 Panels used in RAE 2020. The Information Technology cluster has been highlighted, with the impact pathways for the other clusters greyed out. Figure 2 illustrates the multidisciplinary nature of research impact; multiple journal categories feed into the cluster and the cluster contributes to ICS submitted to a range of RAE panels.

Figure 2: Alluvial diagram linking underpinning research with clusters and panels.
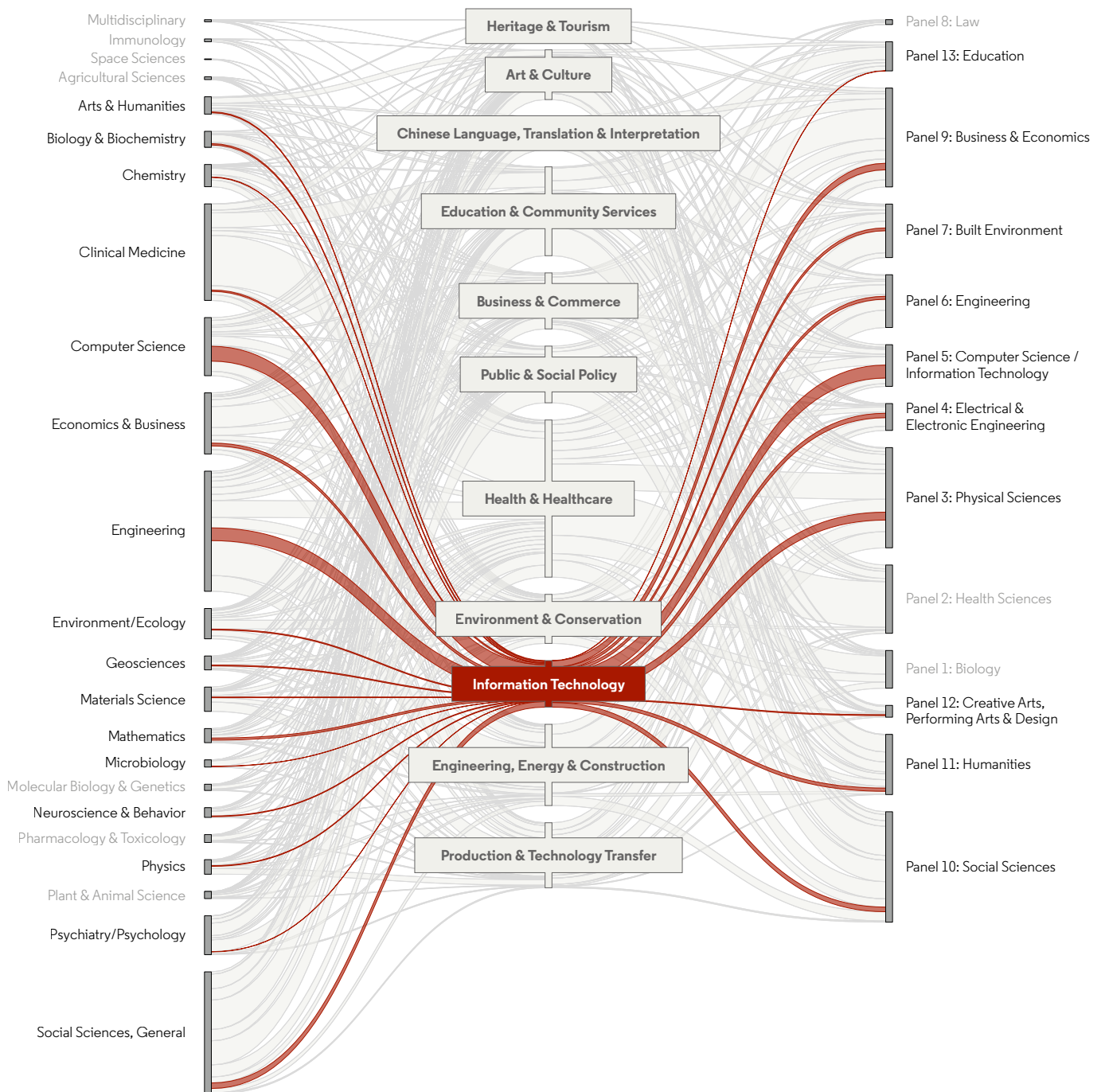
Methodological annex

This synthesised impact report presents a cross-cases analysis of the salient features in 342 impact case studies (ICS) provided by Hong Kong universities as part of the RAE 2020 evaluation. A sequential multi-method approach was employed. The first component involved quantitative topic modelling, followed by directed content analysis. This approach allowed the essence of the impact generated by Hong Kong universities to be captured and synthesised. It is important to note that the analysis and conclusions of these reports are based on the impact as described in the ICS. That is, the authors of this report took the case studies at face value and did not verify or question the narratives provided. A summary of the methodology is given below. For more detailed information on the methodological elements of this study, please see the overarching impact report.

Quantitative topic modelling

Quantitative topic modelling was used to identify overarching topics in the ICS. Topic modelling is a language processing technique applied to document sets to understand the different combinations of words or phrases (topics) that are present. It is a data driven approach, meaning results are not dependent on pre-conceived notions of structure, but are instead derived from the data itself.

Python, Scikit Learn, and Gensim packages were used to implement the topic modelling. Text from section 4 (Details of Impact) from the ICS was normalized (i.e. removal of punctuation and special characters), and domain specific stop-words were removed (i.e. words that are used frequently across the case studies). Various implementations of the topic modelling algorithm were tested, and the Non-negative matrix factorization [NMF] was found to produce the most usable results. After testing multiple models using this algorithm, and manual review by the authors, the number of topics was set to 35 to provide a balance between the breadth of groupings and granularity of topics.

In discussion with UGC, the research team developed an initial taxonomy by grouping similar topics into broader 'clusters'. For example, the topics 'finance', 'accountancy and governance', and 'economics' were grouped into a cluster titled 'business & commerce'. Topic clusters were set at the outset of the analysis to ensure cognitively similar cases were read together, thereby improving the quality of coding, analysis, and impact reports. This classification system then informed the coding and testing of case studies.

Directed content analysis

Qualitative directed content analysis was then used to elucidate the salient characteristics of the impact narratives. This involved an iterative process of examining case studies and developing a code book to categorise their inherent features. The code book was derived from the existing literature and the domain expertise of the authors. It included four overarching categories: a) research, which captured funding source and impetus for research; b) time lags, which captured the elapsed time between the research and its impact; c) mechanisms/channels of impact, which included forms of collaboration and dissemination; and d) impact, which included beneficiary groups (e.g. young people, women, ethnic minorities), location and reach (e.g. Hong Kong, Mainland China, elsewhere), and the nature of impact (e.g. commercial, policy, practice).

Using the cloud based qualitative analysis software, Dedoose, each case study was read, and relevant excerpts were 'tagged' with the relevant codes. Multiple codes and subcodes were attributed to individual case studies. This allowed all case studies that had been tagged with a particular code (e.g. a particular beneficiary group) to be considered as a group. Two of the study's authors undertook the reading and coding (JG and KW). Inter coder reliability was ensured by double coding 10% of the cases (i.e. each author codes the same case study) and through regular coding meetings that were used to compare code applications and adjust the code book as required. The code book was thus a 'living document' that was reviewed and revised iteratively. This process allowed for cross case analysis that was the basis of synthesised impact reports. A code co-occurrence matrix was used to identify where the overarching codes intersect (for example, instances where particular topics are associated with particular beneficiary groups). The properties of the ICS were systematically examined, and evidence was gathered by assigning segments of text to unique codes within the broader coding categories. This process allowed for cross case analysis that formed the basis of this synthesised impact report.

Author information

Kate Williams is Senior Lecturer in Public Policy in the School of Social and Political Sciences at the University of Melbourne and a Visiting Research Fellow at King's College London's Policy Institute. She is currently leading an ESRC Research Grant that compares methods and cultures of research impact evaluation across the U.K., Australia and the U.S. ORCiD: 0000-0002-2882-1068.

Jonathan Grant is founding Director of Different Angles Ltd, a consultancy that focuses on the social impact of universities and research. His main interests are in biomedical and health R&D policy, research impact assessment, the use of research and evidence in policy and decision-taking, and the social purpose of universities in the 21st century. ORCiD: 0000-0002-1646-3486.

Martin Szomszor is founder of Electric Data Solutions and a former Director of ISI. He has expertise in knowledge engineering, machine learning, and natural language processing and was named a 2015 top-50 UK Information Age data leader for his work in creating the REF2014 impact case studies database for the Higher Education Funding Council for England (HEFCE). ORCiD:0000-0003-0347-3527.

Ryan Beardsley is a Senior Consultant for the Academic and Government Consulting Practice at Clarivate™, where his role is to lead projects from inception to completion. He is experienced in bibliometric analysis, which he combines with data science techniques to serve clients from across the globe. Prior to working in research analytics, he spent over a decade as a condensed matter physicist. ORCiD: 0000-0003-4012-6372.

Jonathan Adams is Chief Scientist at the Institute for Scientific Information™ (ISI). He is also a Visiting Professor at King's College London, Policy Institute. In 2017 he was awarded an Honorary D.Sc. by the University of Exeter, for his work in higher education and research policy. ORCiD: 0000-0002-0325-4431.

About Clarivate

Clarivate™ is a global leader in providing solutions to accelerate the pace of innovation. Our bold Mission is to help customers solve some of the world's most complex problems by providing actionable information and insights that reduce the time from new ideas to life-changing inventions in the areas of Academia & Government, Life Sciences & Healthcare, Professional Services and Consumer Goods, Manufacturing & Technology. We help customers discover, protect and commercialize their inventions using our trusted subscription and technology-based solutions coupled with deep domain expertise. For more information, please visit clarivate.com.

About Different Angles

Different Angles Ltd is a consultancy with a distinctive focus on the social impact of universities and research. Founded by Jonathan Grant in 2020, Different Angles provides consultancy, evaluation and training to individuals and institutions in the UK and across the globe. Different Angles clients are a diverse mix of research funders, universities and private innovators & philanthropists. We aim to help clients co-create and co-design innovative, impactful and influential interventions.

About Electric Data Solutions

Electric Data Solutions is an analytics consultancy that uses data science approaches to help research stakeholders understand their unique contribution to the evolving research system. Using a combination of sophisticated data integration techniques, knowledge engineering, and machine learning technology, Electric Data Solutions combines information from multiple sources to generate unique insights into how research is created and utilised. A range of analytical services are provided, including performance benchmarking, horizon scanning and impact assessment, that help customers to make informed strategic decisions in relation to their research investment.

clarivate.com